On the existence of Hamilton cycles with a periodic pattern in a random digraph

Alan Frieze*, Xavier Pérez-Giménez†, Paweł Prałat‡

Abstract

We consider Hamilton cycles in the random digraph $\mathcal{D}_{n,m}$ where the orientation of edges follows a pattern other than the trivial orientation in which the edges are oriented in the same direction as we traverse the cycle. We show that if the orientation forms a periodic pattern, other than the trivial pattern, then approximately half the usual $n \log n$ edges are needed to guarantee the existence of such Hamilton cycles a.a.s.

1 Introduction

The existence of Hamilton cycles is one of the key issues in the study of random graphs. The existence threshold for the random graph process $(\mathcal{G}_{n,m})_{m=0}^{n(n-1)/2}$ was found by Ajtai, Komlós and Szemerédi [1] and the existence threshold for the directed analogue $(\mathcal{D}_{n,m})_{m=0}^{n(n-1)}$ was found by Frieze [5]. (See the next section for definitions of all models mentioned in the introduction as well as the asymptotic notation used.) There is a large literature on this subject and the reader is referred to the bibliography by Frieze [6] for more information.

The result in [5] refers to Hamilton cycles in which the edges are oriented in the same direction round the cycle. Ferber and Long [4] considered Hamilton cycles with prescribed edge orientations. They proved that if C_n is an n-cycle with an arbitrary edge orientation then if $np \gg (\log\log n)\log n$ then $\mathcal{D}_{n,p}$ contains a copy of C_n a.a.s. (In fact, they proved more than this, in that they proved the existence of many copies.) They conjectured that, if $np = \log n + \omega(1)$, then this is sufficient for $\mathcal{D}_{n,p}$ to contain a copy of C_n a.a.s. Montgomery [8] has recently announced a proof of this conjecture.

In this paper we consider Hamilton cycles with a periodic pattern of edge orientations and show that we need approximately half as many random edges to guarantee the existence of such Hamilton cycles a.a.s.

The paper is structured as follows. In the next section, we introduce all the definitions and state the main results. Section 3 outlines the proof. Section 4 provides structural results, and Section 5 uses these results to construct the claimed cycles.

*Department of Mathematical Sciences, Carnegie Mellon University, Pittsburgh PA, USA, 15213. Research supported in part by NSF grant DMS1661063.

†Department of Mathematics, University of Nebraska-Lincoln, Lincoln NE, USA, 68588. Research supported in part by Simons Foundation Grant #587019.

‡Department of Mathematics, Ryerson University, Toronto, ON, Canada M5B 2K3. Research supported in part by NSERC Discovery Grant.

2 Definitions and Main Results

2.1 Random Digraphs

In this paper we present results obtained for the *random directed graph* $\mathcal{D}_{n,p}$. More precisely, $\mathcal{D}_{n,p}$ is a distribution over the class of graphs with vertex set $[n] := \{1, \dots, n\}$ in which every ordered pair ij, $i, j \in [n]$ and $i \neq j$, appears independently as an arc in $\mathcal{D}_{n,p}$ with probability p. Note that $p = p(n)$ may (and usually does) tend to zero as n tends to infinity. We say that $\mathcal{D}_{n,p}$ has some property *asymptotically almost surely* (or a.a.s.) if the probability that $\mathcal{D}_{n,p}$ has this property tends to 1 as n goes to infinity.

As it is done in the theory of (undirected) random graphs, we slightly abuse the notation and for some natural number $m = m(n) \in \mathbb{N}$ use $\mathcal{D}_{n,m}$ to denote a directed graph on n vertices and precisely m arcs taken uniformly at random from the family of directed graphs on n vertices and m arcs. Alternatively, $\mathcal{D}_{n,m}$ can be constructed during the *random digraph process*. Indeed, one may consider a sequence of digraphs $(\mathcal{D}_{n,m})_{m=0}^{n(n-1)}$ with common vertex set $[n]$ in which $\mathcal{D}_{n,0}$ is the empty digraph and $\mathcal{D}_{n,m+1}$ is obtained from $\mathcal{D}_{n,m}$ by adding a random arc that is not present in $\mathcal{D}_{n,m}$. The process ends once $\mathcal{D}_{n,n(n-1)}$, the complete digraph, is reached.

2.2 Asymptotic Notation and Convention

Given two nonnegative functions $f = f(n)$ and $g = g(n)$, we will write $f = O(g)$ if there exists an absolute constant c such that $f(n) \leq cg(n)$ for all $n \in \mathbb{N}$, $f = \Omega(g)$ if $g = O(f)$, $f = \Theta(g)$ if $f = O(g)$ and $f = \Omega(g)$, and we write $f = o(g)$ or $f \ll g$ if the limit $\lim_{n\to\infty} f(n)/g(n) = 0$. In addition, we write $f = \omega(g)$ or $f \gg g$ if $g = o(f)$. We also will write $f \sim g$ if $f = (1 + o(1))g$.

Through the paper, all logarithms with no subscript denoting the base will be taken to be natural. Moreover, as typical in the field of random graphs, for expressions that clearly have to be an integer, we round up or down but do not specify which, as long as this rounding does not affect the argument.

2.3 Patterns

Given a cycle $C = (v_1, \dots, v_n, v_1)$ of length $n \geq 1$ (where C is a loop for $n = 1$ and a double arc for $n = 2$), we are interested in describing periodic orientations of the edges of C. A *pattern* of length $k \geq 1$ is a k-tuple $\pi = (\pi_1, \dots, \pi_k)$ where $\pi_i \in \{\rightarrow, \leftarrow\}$. We say that an orientation of C *follows* pattern π if $k \mid n$ and each edge v_iv_{i+1} for $i \in [n]$ is oriented as $\overrightarrow{v_iv_{i+1}}$ if $\pi_i = \rightarrow$ or as $\overleftarrow{v_iv_{i+1}}$ otherwise. (Here $v_{n+1} = v_1$, and the indices of π_i should be taken modulo k.) We identify a pattern π with the only oriented cycle C_π of length k which follows π, and consider two patterns π and π' to be *equivalent* if their corresponding oriented cycles C_π and $C_{\pi'}$ are isomorphic digraphs. In other words, two patterns are equivalent if they can be obtained from one another by applying a cyclic rotation of the entries and/or a reflection (that is, reversing *both* the order of the entries and the direction of the arrows). A pattern is *primitive* if it is not a concatenation of two or more consecutive copies of a shorter pattern. We call an oriented cycle C in a digraph D a *π-Hamilton cycle* (or *π-HC*

for short) if C spans D and follows pattern π. Clearly, for two equivalent patterns π and π', a digraph D contains a π-HC if and only if it contains a π'-HC. Moreover, if π is a primitive pattern of length k and π' is the concatenation of ℓ copies of π, then a digraph on n vertices has a π'-HC if and only if it has a π-HC and $\ell k \mid n$. In view of that, we can restrict our attention to the analysis of primitive patterns π. All patterns of length 1 are equivalent to $(\rightarrow)$, which we call the *trivial* pattern. Likewise, all primitive patterns of length 2 are equivalent to $(\rightarrow, \leftarrow)$. We refer to $(\rightarrow, \leftarrow)$ as the *alternating* pattern, so a $(\rightarrow, \leftarrow)$-HC is simply called an alternating Hamilton cycle. Primitive patterns of length $k \geq 3$ are defined to be *non-alternating*.

Remark 1. *While our main results concern only primitive patterns, in some parts of the argument it will be convenient to consider the non-primitive pattern $(\rightarrow, \leftarrow, \rightarrow, \leftarrow)$ as well. This pattern will be useful when investigating the existence of alternating Hamilton cycles. Indeed, when the number of vertices n is divisible by 4, a Hamilton cycle follows $(\rightarrow, \leftarrow)$ if and only if it follows $(\rightarrow, \leftarrow, \rightarrow, \leftarrow)$. In view of this, we can avoid some of the technical difficulties that arise in the analysis of patterns of length 2 by simply redefining the alternating pattern to be $(\rightarrow, \leftarrow, \rightarrow, \leftarrow)$ when $4 \mid n$. The case when $n \equiv 2 \mod 4$ can be reduced to the previous case with a slight modification of the argument.*

2.4 Results

We are now ready to state our main results.

Theorem 2. *Let $\pi = (\rightarrow, \leftarrow)$ be the alternating pattern. Consider the random digraph process $(\mathcal{D}_{n,m})_{m=0}^{n(n-1)}$ (restricted to even n). The following property holds a.a.s.: $\mathcal{D}_{n,m}$ contains a π-HC the first time all vertices have in-degree at least 2 or out-degree at least 2.*

From this result, we immediately and easily establish the sharp threshold for the existence of the alternating Hamilton cycle, and obtain the limiting probability at the critical window. In particular, the following corollary holds.

Corollary 3. *If $p = \frac{\log n + 2\log\log n + c}{2n}$ and n is even, then*

$$\mathbf{Pr}(\mathcal{D}_{n,p} \text{ contains an alternating Hamilton cycle}) \sim e^{-e^{-c}/4}.$$

We obtain similar results for all primitive patterns of order $k \geq 3$.

Theorem 4. *For any fixed $k \geq 3$, let $\pi = (\pi_1, \ldots, \pi_k)$ be a primitive pattern. Consider the random digraph process $(\mathcal{D}_{n,m})_{m=0}^{n(n-1)}$ (restricted to integers n that are divisible by k). The following property holds a.a.s.: $\mathcal{D}_{n,m}$ contains a π-HC the first time all vertices have total degree at least 2.*

Corollary 5. *Let $\pi = (\pi_1, \ldots, \pi_k)$ be a primitive pattern of length $k \geq 3$. If $p = \frac{\log n + \log\log n + c}{2n}$ and $k \mid n$, then*

$$\mathbf{Pr}(\mathcal{D}_{n,p} \text{ has } \pi\text{-HC}) \sim e^{-e^{-c}}.$$

Remark 6. *Let us point out that there are different coefficients in the $\log\log n$ terms in Corollaries 3 and 5. It follows from the fact that having a vertex with both in-degree and out-degree equal to 1 prevents there being a Hamilton cycle with the alternating pattern, but it does* not *prevent any other pattern. See Lemma 10 to see how this affects the probability threshold.*

3 Idea of the Proof

We start with a naïve description of the argument. Let $k \geq 1$ be any fixed natural number, and let π be any pattern of length k. Suppose that we wish to find a π-Hamilton cycle in a given random digraph on n vertices, where n is divisible by k. Here is one simple approach: first partition the n vertices into k bins, $B_1, \ldots, B_k$, so that each bin receives exactly n/k vertices (this is done arbitrarily, before examining the edges); then build a cycle $(v_1, \ldots, v_n, v_1)$, where each $v_i \in B_i$ and each $v_i v_{i+1}$ is oriented as π_i. (Here the indices of B_i and π_i should be interpreted modulo k). This gives us the desired π-HC. Clearly, for the above construction to succeed, every vertex v in bin B_i should have some neighbours $w \in B_{i-1}$ and $w' \in B_{i+1}$ such that the edges vw and vw' are appropriately oriented. Unfortunately, it is easy to see that for the $\mathcal{D}_{n,p}$ model the above property fails a.a.s. when $k \geq 2$ and $p \sim \log n/(2n)$ (which is the sharp threshold for the degree constraints given in Theorems 2 and 4). This is due to the presence of vertices v such that v and all its neighbours belong to the same bin. To overcome this obstacle, in this section we will analyze a simpler model of random digraphs for which the above naïve construction of a π-HC works. Later in Section 5, we will show how to modify $\mathcal{D}_{n,p}$ (by moving a few vertices to different bins and contracting some short paths) so that the resulting random graph can be analyzed by means of this simpler model.

We consider a model that generalizes the $D_{s\text{-in},t\text{-out}}$ model of [2] as follows. Let $S = (s_{i,j})_{1 \leq i,j \leq k}$ and $T = (t_{i,j})_{1 \leq i,j \leq k}$ be two $k \times k$ matrices with entries in $\mathbb{N} \cup \{0\}$. We construct the random digraph $D_{S\text{-in},T\text{-out}}$ as follows. We have n vertices placed in $k \geq 1$ bins as explained above. For each vertex v in bin B_i, we add $s_{i,j}$ distinct in-arcs to v (that is, with head vertex v) and tail vertices chosen uniformly at random (u.a.r.) from $B_j \setminus \{v\}$. In other words, each of the $\binom{n/k}{s_{i,j}}$ choices (or $\binom{n/k-1}{s_{i,j}}$ choices if $i = j$) is equally likely. Similarly, for each vertex v in bin B_i, we add $t_{i,j}$ distinct out-arcs from v (that is, with tail vertex v) and head vertices chosen u.a.r. from $B_j \setminus \{v\}$. Let us note that $D_{S\text{-in},T\text{-out}}$ is, in general, a multi-digraph since an arc vw may be generated once as an out-arc from v and once as an in-arc to w.

We start with the following observation.

Lemma 7. *Let $S = T = \begin{pmatrix} 0 & 2 \\ 2 & 0 \end{pmatrix}$. Then a.a.s. $D_{S\text{-in},T\text{-out}}$ has one perfect matching directed from B_1 to B_2 and one directed from B_2 to B_1.*

Proof. Walkup [9] (see Theorem 17.6 in [7]) proved that the following model of a random bipartite graph has a perfect matching a.a.s: given disjoint sets A, B of size m, each vertex of A chooses 2 random neighbours in B and each vertex of B chooses 2 random neighbours in A. This is a random graph with $4m$ edges. The random graph $D_{S\text{-in},T\text{-out}}$ consists of two independent copies of Walkup's graph, with orientations preserved, and so the conclusion follows immediately. □

With one additional ingredient, namely, the fact that $D_{2\text{-in},2\text{-out}}$ a.a.s. has a directed Hamilton cycle, we get the following useful lemma. For technical reasons we disregard the case $k = 2$, and restrict our attention to patterns of length $k \geq 3$ (see Remark 9 below).

Moreover, we do not assume patterns to be primitive so, in particular, the lemma applies to $\pi = (\rightarrow, \leftarrow, \rightarrow, \leftarrow)$.

Lemma 8. *Let $k \geq 3$ be any natural number and let $\pi = (\pi_1, \ldots, \pi_k)$ be any pattern. Let $S = (s_{i,j})_{1 \leq i,j \leq k}$ and $T = (t_{i,j})_{1 \leq i,j \leq k}$ be $k \times k$ matrices with $s_{i,i+1} = s_{i+1,i} = t_{i,i+1} = t_{i+1,i} = 2$. Then a.a.s. $D_{S\text{-}in,T\text{-}out}$ has a Hamilton cycle $(v_1, \ldots, v_n)$ with $v_i \in B_j$ when $i \equiv j \mod k$ and with each $v_i v_{i+1}$ oriented as π_i.*

Proof. Using Lemma 7, we find a family of perfect matchings between B_i and B_{i+1} directed as π_i for each $i = 1, \ldots, k-1$. This gives a partition of the vertex set into sets that induce paths of length $k-1$. Each of these paths $P = (u_1, \ldots, u_k)$ has $u_i \in B_i$ and $u_i u_{i+1}$ oriented as π_i. We contract each path into a single vertex. Without loss of generality, we may assume that $\pi_k = \rightarrow$; otherwise, a symmetric argument can be applied. It suffices to find a directed Hamilton cycle on the contracted paths, so that each edge is oriented from the B_k end of one path to the B_1 end of the next one. In order to do this, given each path $P = (u_1, \ldots, u_k)$, note there are $t_{k,1} = 2$ out-arcs from u_k to a random vertex in B_1 and $s_{1,k} = 2$ in-arcs to u_1 from a random vertex in B_k. Keeping only those arcs, we get a $D_{2\text{-in},2\text{-out}}$ graph on the vertices associated with the contracted paths. It was shown by Cooper and Frieze [3] that $D_{2\text{-in},2\text{-out}}$ a.a.s. has a directed Hamilton cycle. Un-contracting the paths gives us the desired π-Hamilton cycle of the original graph. □

Remark 9. *The proof of the lemma does not work 'as is' for $k = 2$ and $\pi = (\rightarrow, \leftarrow)$, since the arcs between B_1 and B_2 used in the last step to complete a Hamilton cycle may have already been used in the first perfect matching between these same bins. With some extra work, one can show that the statement still holds in that scenario if we allow $S = T = \begin{pmatrix} 0 & 4 \\ 4 & 0 \end{pmatrix}$. We chose not to include this case in the statement, since it is not needed in our main argument.*

Lemma 8 will be useful in analyzing the random digraph process. However, the argument will be more delicate and the above proof strategy needs to be amended in order to deal with vertices of low degree that occur in the original process but not in the $D_{S\text{-in},T\text{-out}}$ one. These vertices will be "buried" inside paths after contracting them to "fat" vertices.

4 Structural Ingredients

4.1 Bin partition

Throughout this section, π is a fixed primitive nontrivial pattern. We consider two different scenarios: either $\pi = (\rightarrow, \leftarrow)$ (alternating case) or π is a primitive pattern $(\pi_1, \ldots, \pi_k)$ of length $k \geq 3$ (non-alternating case). We will analyze several models of random digraphs with common vertex set $[n]$. In the non-alternating case, we assume that the length of the pattern k divides n, and partition the vertices into k bins $B_1, \ldots, B_k$ in an equitable manner (that is, every bin receives exactly n/k vertices). The alternating case is slightly different. Here we assume that n is even, but partition the vertices into 4 bins $B_1, \ldots, B_4$, so that bins B_1, B_2 receive $\lceil n/4 \rceil$ vertices and bins B_3, B_4 receive $\lfloor n/4 \rfloor$ vertices. By setting $k = 4$, the alternating case can be intuitively regarded as if $\pi = (\rightarrow, \leftarrow, \rightarrow, \leftarrow)$ instead of $(\rightarrow, \leftarrow)$, with

the weaker requirement that n is even but not necessarily divisible by 4 (see Remark 1). In either case, the partition of the vertex set into bins is done arbitrarily before examining the arcs of any of the random digraphs.

4.2 Necessary conditions

We now describe what will turn out to be the main obstruction to the existence of a π-HC in $\mathcal{D}_{n,p}$. In the non-alternating case (that is, π primitive of length $k \geq 3$), let $\mathcal{A}$ denote the event that $\mathcal{D}_{n,p}$ has no vertices of total degree less than 2. In the alternating case, we require a stronger condition, and define $\mathcal{A}$ to be the event that every vertex in $\mathcal{D}_{n,p}$ has in-degree at least 2 or out-degree at least 2. Note that in either case $\mathcal{A}$ is trivially a necessary condition for the existence of π-HC.

The following result is a standard exercise in the field of random graphs, so we just sketch the main steps of the proof.

Lemma 10. *Let $c \in \mathbb{R}$ be any fixed constant.*

1. *For $\pi = (\rightarrow, \leftarrow)$ and with $p_{alt} = p_{alt}(n) = \frac{\log n + 2\log\log n + c}{2n}$*

$$\mathbf{Pr}(\mathcal{D}_{n,p_{alt}} \text{ satisfies } \mathcal{A}) \sim e^{-e^{-c}/4}.$$

2. *For a primitive pattern π of length $k \geq 3$ and with $p_{non\text{-}alt} = p_{non\text{-}alt}(n) = \frac{\log n + \log\log n + c}{2n}$,*

$$\mathbf{Pr}(\mathcal{D}_{n,p_{non\text{-}alt}} \text{ satisfies } \mathcal{A}) \sim e^{-e^{-c}}.$$

Proof (sketch). Let us first consider the case $\pi = (\rightarrow, \leftarrow)$. Let X denote the number of vertices in $\mathcal{D}_{n,p_{\text{alt}}}$ with in-degree 1 and out-degree 1. Easy computations show that

$$\mathbf{E}X \sim n^3 p_{\text{alt}}^2 (1-p_{\text{alt}})^{2n} \sim \frac{n\log^2 n}{4} \exp\Big(-(\log n + 2\log\log n + c)\Big) = e^{-c}/4.$$

Similarly, the ith factorial moment satisfies $\mathbf{E}[(X)_i] \sim (e^{-c}/4)^i$ for each fixed $i \in \mathbb{N}$. As a result, it follows that X is asymptotically Poisson with mean $\lambda := e^{-c}/4$. In particular, $\mathbf{Pr}(X = 0) \sim e^{-\lambda}$. Moreover, the expected number of vertices of total degree at most 1 is equal to $O(1/\log n) = o(1)$, and so

$$\mathbf{Pr}(\mathcal{A}) = \mathbf{Pr}(X = 0) + o(1) \sim e^{-e^{-c}/4}.$$

The non-alternating case is very similar. This time we need to investigate random variable Y, the number of vertices in $\mathcal{D}_{n,p_{\text{non-alt}}}$ with total degree 1. We get that

$$\mathbf{E}Y \sim n^2 (2p_{\text{non-alt}})(1-p_{\text{non-alt}})^{2n} \sim n\log n \exp\Big(-(\log n + \log\log n + c)\Big) = e^{-c},$$

and so

$$\mathbf{Pr}(\mathcal{A}) = \mathbf{Pr}(Y = 0) + o(1) \sim e^{-e^{-c}}.$$

The details are left for the reader. □

4.3 In- and out-arcs

For each $p = p(n) \in [0,1]$, we can view $\mathcal{D}_{n,p}$ as the union of two independent copies of $\mathcal{D}_{n,p'}$, denoted $\mathcal{D}^{\text{in}}_{n,p'}$ and $\mathcal{D}^{\text{out}}_{n,p'}$, where $p' = p'(n)$ is chosen such that

$$2p' - p'^2 = p.$$

This is a standard, simple but useful, observation called *two-round exposure* or *sprinkling*. The arcs in $\mathcal{D}^{\text{in}}_{n,p'}$ receive label"in", and we will call them *in-arcs*. Likewise, the arcs in $\mathcal{D}^{\text{out}}_{n,p'}$ are labeled "out", and will be called *out-arcs*. Note that an arc $\overrightarrow{vw}$ may appear in both $\mathcal{D}^{\text{in}}_{n,p'}$ and $\mathcal{D}^{\text{out}}_{n,p'}$, which creates a parallel pair of arcs, one in-arc $\overrightarrow{vw}$ and one out-arc $\overrightarrow{vw}$. This is useful for the purpose of building a Hamilton cycle, since we may use either copy of $\overrightarrow{vw}$ for the cycle. By forgetting the labels and merging parallel pairs into single arcs, we recover the usual distribution of $\mathcal{D}_{n,p}$. Hence, we regard $\mathcal{D}_{n,p}$ as a simple digraph, where each arc has label "in", "out", or both.

Furthermore, we wish to build the standard random digraph process $(\mathcal{D}_{n,p})_{0\le p\le 1}$ in a way that is compatible with the arc labels. To do so, we introduce a sequence $(X^{\text{in}}_{v,w}, X^{\text{out}}_{v,w})_{v,w\in[n],v\neq w}$ of i.i.d. random variables uniformly distributed in $[0,1]$. Then, for each pair of different vertices v, w, we include arc $\overrightarrow{vw}$ in $\mathcal{D}^{\text{in}}_{n,p'}$ (or in $\mathcal{D}^{\text{out}}_{n,p'}$) if $X^{\text{in}}_{v,w} \le p'$ (or, respectively, $X^{\text{out}}_{v,w} \le p'$). Hence, setting $\mathcal{D}_{n,p} = \mathcal{D}^{\text{in}}_{n,p'} \cup \mathcal{D}^{\text{out}}_{n,p'}$ (and merging parallel pairs into single arcs), we obtain three random processes $(\mathcal{D}_{n,p})_{0\le p\le 1}$, $(\mathcal{D}^{\text{in}}_{n,p'})_{0\le p'\le 1}$ and $(\mathcal{D}^{\text{out}}_{n,p'})_{0\le p'\le 1}$ with the usual couplings: $\mathcal{D}_{n,p_1} \subseteq \mathcal{D}_{n,p_2}$, $\mathcal{D}^{\text{in}}_{n,p'_1} \subseteq \mathcal{D}^{\text{in}}_{n,p'_2}$ and $\mathcal{D}^{\text{out}}_{n,p'_1} \subseteq \mathcal{D}^{\text{out}}_{n,p'_2}$ for $0 \le p_1 \le p_2 \le 1$. Note that, if $X^{\text{in}}_{v,w} = p'_1 < p'_2 = X^{\text{out}}_{v,w}$, we say that arc $\overrightarrow{vw}$ appears in $(\mathcal{D}_{n,p})_{0\le p\le 1}$ at time p_1 with label "in" and gets a second label "out" at time p_2. However, if we choose to ignore labels, then arc $\overrightarrow{vw}$ simple appears in $(\mathcal{D}_{n,p})_{0\le p\le 1}$ at time p_1.

In the context of the random digraph process $(\mathcal{D}_{n,p})_{0\le p\le 1}$, we consider random variable

$$p_* = \min\{p \in [0,1] : \mathcal{D}_{n,p} \text{ satisfies } \mathcal{A}\}.$$

Let $\omega := \log\log\log n = o(\log\log n)$ and let

$$p_\pm = \begin{cases} \dfrac{\log n + 2\log\log n \pm \omega}{2n} & \text{if } \pi \text{ is the alternating pattern } (k=2), \\[2ex] \dfrac{\log n + \log\log n \pm \omega}{2n} & \text{otherwise } (k \ge 3). \end{cases} \tag{1}$$

In particular, by Lemma 10, a.a.s. $p_- < p_* < p_+$.

In the remainder of this section, we will analyze the two random digraph processes $(\mathcal{D}^{\text{in}}_{n,p'})_{p'_-\le p'\le p'_+}$ and $(\mathcal{D}^{\text{out}}_{n,p'})_{p'_-\le p'\le p'_+}$, where $2p'_\pm - {p'_\pm}^2 = p_\pm$. These two processes, in turn, determine the process $(\mathcal{D}_{n,p})_{p_-\le p\le p_+}$, as explained above. For each "time" $p' \in [p'_-, p'_+]$, we associate out-arcs (that is, arcs in $\mathcal{D}^{\text{out}}_{n,p'}$) to their tail vertex and in-arcs to their head vertex. More precisely, if $\overrightarrow{vw}$ is an out-arc with tail at v and head at w, then we say that it is *visible* from v and *invisible* from w. Similarly, if $\overrightarrow{vw}$ is an in-arc, then it is *invisible* from v and *visible* from w. Later in the argument, we will expose the endpoint of each arc from which the arc is visible and leave the other endpoint random. This trick will be useful to emulate the $D_{S\text{-in},T\text{-out}}$ model.

4.4 Useful properties

Recall that vertices are equitably partitioned into bins $B_1, \dots, B_k$ before exposing any random arcs. (In the alternating case, $k = 4$ and $|B_1| = |B_2| = \lceil n/4 \rceil$ while $|B_3| = |B_4| = \lfloor n/4 \rfloor$.) We use this partition to classify vertices into three different types (good, bad, or dangerous):

- We say a vertex v is *good* if, for every $i \in [k]$, v has at least $4k+2$ visible out-arcs (that is, out-arcs visible from v) in $\mathcal{D}^{\text{out}}_{n,p'_-}$ whose head vertex is in bin B_i and also at least $4k+2$ visible in-arcs in $\mathcal{D}^{\text{in}}_{n,p'_-}$ whose tail is in B_i. (Note that this property is required for each bin, including the bin vertex v belongs to.)
- We say that v is *bad* if it is not good, but its total degree in $\mathcal{D}_{n,p_-} = \mathcal{D}^{\text{in}}_{n,p'_-} \cup \mathcal{D}^{\text{out}}_{n,p'_-}$ (including all arcs that are incident with v, visible and invisible, and ignoring labels) is at least $4k+3$.
- The remaining vertices, which have total degree at most $4k+2$, are called *dangerous*.

The above classification of vertices is based solely on the two digraphs $\mathcal{D}^{\text{in}}_{n,p'_-}$ and $\mathcal{D}^{\text{out}}_{n,p'_-}$. It remains invariant throughout $(\mathcal{D}^{\text{in}}_{n,p'})_{p'_- \le p' \le p'_+}$, $(\mathcal{D}^{\text{out}}_{n,p'})_{p'_- \le p' \le p'_+}$, and hence $(\mathcal{D}_{n,p})_{p_- \le p \le p_+}$. In other words, if a vertex is good/bad/dangerous at time p_- it stays of this type during any time $p \in [p_-, p_+]$.

Given a digraph $\mathcal{D}$, the $\mathcal{D}$-*distance* between two vertices of $\mathcal{D}$ denotes the usual graph distance in the underlying undirected graph. Let $\mathcal{H}$ (for handsome) be the family of all digraphs $\mathcal{D}$ with $\mathcal{D}_{n,p_-} \subseteq \mathcal{D} \subseteq \mathcal{D}_{n,p_+}$ satisfying:

H1: Every vertex in $\mathcal{D}$ has fewer than $4k$ vertices that are not good within $\mathcal{D}$-distance $10k$.

H2: Every dangerous vertex in $\mathcal{D}$ has only good vertices within $\mathcal{D}$-distance $10k$.

H3: Every cycle (with any orientation, any labels on the associated arcs, and including cycles of length 2) in $\mathcal{D}$ of length at most $10k$ has only good vertices within $\mathcal{D}$-distance $10k$.

Note that distances in **H1**–**H3** are measured with respect to digraph $\mathcal{D}$, whereas the definitions of being good, bad, and dangerous depend only on arcs that are already present in $\mathcal{D}_{n,p_-} \subseteq \mathcal{D} \subseteq \mathcal{D}_{n,p_+}$. As a result, $\mathcal{H}$ is a monotone decreasing family of digraphs.

We will show now that random directed graphs are typically handsome.

Lemma 11. *A.a.s. $\mathcal{D}_{n,p} \in \mathcal{H}$ for all $p_- \le p \le p_+$.*

Proof. As mentioned above, **H1**–**H3** are monotone decreasing properties, so it suffices to show that a.a.s. $\mathcal{H}$ holds for $p = p_+$. We will investigate these three events in turn.

H1: Let $M = 40k^2$. Then,

$$\begin{aligned}
&\mathbf{Pr}(\exists \text{ a vertex of } \mathcal{D}_{n,p_+} \text{ not satisfying } \mathbf{H1}) \\
&\leq n \sum_{\ell=4k}^{M} \binom{n}{\ell} (\ell+1)^{\ell-1} (2p_+)^{\ell} \binom{\ell}{4k} \left(2k \sum_{i=0}^{4k+1} \binom{\lceil n/k \rceil}{i} p_-^i (1-p_-)^{\lfloor n/k \rfloor - M - 1 - i} \right)^{4k} \\
&\leq n \sum_{\ell=4k}^{M} \left(\frac{ne}{\ell}\right)^{\ell} \left(\frac{2(\ell+1)\log n}{n}\right)^{\ell} \\
&\qquad \times \left(\frac{e\ell}{4k} (2k) 2 (n/k)^{4k+1} (\log n/n)^{4k+1} \exp\left(-(1+o(1)) \frac{\log n}{2n} \cdot \frac{n}{k} \right) \right)^{4k} \\
&\leq n \sum_{\ell=4k}^{M} (10 \log n)^{\ell} \left(\log^{4k+1} n \cdot n^{-1/(2k)+o(1)} \right)^{4k} = n^{1-2+o(1)} = o(1).
\end{aligned}$$

Indeed, there are n choices for a vertex v and then there will be an additional ℓ vertices that make up the paths, where $4k \leq \ell \leq M = (4k) \times (10k)$. We choose these vertices in at most $\binom{n}{\ell}$ ways and then choose a spanning tree that is contained in the union of the paths in at most $(\ell+1)^{\ell-1}$ ways, the number of labelled trees on $\ell+1$ vertices. The factor $(2p_+)^{\ell}$ bounds the probability the arcs of the tree exist, *ignoring orientation* gives us the 2. We then choose $4k$ vertices that are not good and multiply by an upper bound for the probability that these vertices have few $\mathcal{D}_{n,p_-}$ in- or out-neighbours in some of the k bins outside of the set of $\ell + 1 \leq M + 1$ vertices chosen so far.

H2: The probability that there is a dangerous vertex v and a not good vertex w such that $\text{dist}(v,w) = \ell \leq 10k$ can be bounded from above by

$$\begin{aligned}
&n \sum_{\ell=1}^{10k} n^{\ell} (2p_+)^{\ell} \left(2k \sum_{i=0}^{4k+1} \binom{\lceil n/k \rceil}{i} p_-^i (1-p_-)^{\lfloor n/k \rfloor - \ell - 1 - i} \right) \\
&\qquad \times \left(\sum_{j=0}^{4k+1} \binom{n}{j} (2p_-)^j (1 - 2p_- + p_-^2)^{n-2-j} \right) \\
&\leq n \sum_{\ell=1}^{10k} \Big((1+o(1)) \log n \Big)^{\ell} n^{-1/(2k)+o(1)} n^{-1+o(1)} = n^{-1/(2k)+o(1)} = o(1).
\end{aligned}$$

Indeed, there are n choices for a dangerous vertex v and then at most n^{ℓ} choices for additional vertices that form a path of length ℓ reaching a vertex w that is not good. The next term is an upper bound for the probability that w is not good. The last term is an upper bound for the probability that v is dangerous. Note that v already has one neighbour on the path that is already identified; hence, $j \leq 4k+1$, not $4k+2$.

H3: If there is a cycle not satisfying **H3**, then there is a set of $2 \leq \ell \leq M = 10k + 10k = 20k$ vertices containing at least ℓ arcs and at least one not good vertex. The probability of

this event occurring in $\mathcal{D}_{n,p}$ can be bounded from above by

$$\sum_{\ell=2}^{M}\binom{n}{\ell}\binom{\ell^2}{\ell}(2p_+)^\ell\ell(2k)\sum_{i=0}^{4k+1}\binom{\lceil n/k\rceil}{i}p_-^i(1-p_-)^{\lfloor n/k\rfloor-M-i}$$
$$\leq 40k^2\sum_{\ell=2}^{M}\left(\frac{ne}{\ell}\right)^\ell\left(\frac{\ell^2e}{\ell}\right)^\ell\left(\frac{2\log n}{n}\right)^\ell n^{-1/(2k)+o(1)}=n^{-1/(2k)+o(1)}=o(1).$$

The proof of the lemma is finished. □

Given a digraph $\mathcal{D}_{n,p_-}\subseteq\mathcal{D}\subseteq\mathcal{D}_{n,p_+}$ and a vertex v of $\mathcal{D}$, the neighbourhood $N_{\mathcal{D}}(v)=N(v)$ of v is the set of all vertices w that are joined to v by an arc (in- or out-arc, visible or invisible from v). We now give some useful deterministic consequences of properties **H1**–**H3**.

Lemma 12. *Let $\mathcal{D}$ be any digraph in $\mathcal{H}$, and let $\mathscr{P}_0$ be a pairwise vertex-disjoint collection of paths, where each path $P\in\mathscr{P}_0$ has length at most $6k+2$ and contains exactly one non-good vertex (that is, bad or dangerous). Then*

H4 *For every dangerous vertex v, all the vertices in $N(v)$ are good and none of them is contained in a path in $\mathscr{P}_0$.*

H5 *For every vertex v, all but at most $4k$ vertices in $N(v)$ are good and are not contained in a path in $\mathscr{P}_0$.*

Proof. Pick any vertex v. **H3** implies that no path $P\in\mathscr{P}_0$ intersects more than one vertex in $N(v)$, since otherwise we would create a cycle of length at most $10k$ with at least one non-good vertex within distance $10k$. Then, by **H1**, we immediately get **H5**. Moreover, if v is dangerous, then **H2** implies **H4**. □

5 Cycle Construction

In this section we will prove Theorems 2 and 4 by showing that a.a.s. $\mathcal{D}_{n,p_*}$ contains a π-Hamilton cycle, where π is either the alternating pattern $(\rightarrow,\leftarrow)$ or a primitive pattern $(\pi_1,\ldots,\pi_k)$ of length $k\geq 3$. For notational convenience, in the alternating case we set $k=4$ and redefine π to be $(\rightarrow,\leftarrow,\rightarrow,\leftarrow)$, even though n is only assumed to be even but not necessarily divisible by 4 (see Remark 1). In the non-alternating case, we always require that $k\mid n$. As in Section 4, the vertex set $[n]$ is equitably partitioned into $k\geq 3$ bins $B_1,\ldots,B_k$ of size $|B_i|=n/k$ (in the non-alternating case) or $|B_1|=|B_2|=\lceil n/4\rceil$ and $|B_3|=|B_4|=\lfloor n/4\rfloor$ (in the alternating case). Recall that this partition is done before exposing any random arcs.

We will analyze the random processes $(\mathcal{D}^{\text{in}}_{n,p'})_{p'_-\leq p'\leq p'_+}$, $(\mathcal{D}^{\text{out}}_{n,p'})_{p'_-\leq p'\leq p'_+}$ and $(\mathcal{D}_{n,p})_{p_-\leq p\leq p_+}$ introduced in Section 4, where $\mathcal{D}_{n,p}=\mathcal{D}^{\text{in}}_{n,p'}\cup\mathcal{D}^{\text{out}}_{n,p'}$ for $2p'-{p'}^2=p$ (merging parallel arcs into single arcs) and where p_-,p_+ are defined as in (1). In view of Lemmas 10 and 11, we will assume that $p_-<p_*<p_+$ and that event $\mathcal{H}$ holds in $\mathcal{D}_{n,p}$ for all $p_-\leq p\leq p_+$. Otherwise our construction will simply fail, but this occurs with probability $o(1)$. Note that we do not

condition on these events, since this would destroy the probability space. Instead, we will expose some partial information about $(\mathcal{D}_{n,p})_{p_- \le p \le p_+}$ and assume it does not contradict our assumptions above, while we leave all the unexposed information random.

Step 1 – digraph $\mathcal{D}_*$: We first consider digraphs $\mathcal{D}^{\text{in}}_{n,p'_-}$ and $\mathcal{D}^{\text{out}}_{n,p'_-}$ (or equivalently digraph $\mathcal{D}_{n,p_-}$ at time p_-). Recall that an in-arc $\overrightarrow{vw}$ of $\mathcal{D}^{\text{in}}_{n,p'_-}$ is visible from w, while an out-arc $\overrightarrow{vw}$ of $\mathcal{D}^{\text{out}}_{n,p'_-}$ is visible from v. For every vertex v and for every bin B_i, we expose the number $d_i^-(v)$ of in-arcs in $\mathcal{D}^{\text{in}}_{n,p'_-}$ with head at v and tail at some vertex in B_i and similarly the number $d_i^+(v)$ of out-arcs in $\mathcal{D}^{\text{out}}_{n,p'_-}$ with tail at v and head at some vertex in B_i. (In other words, we reveal the number of arcs between v and B_i in each orientation and which are visible from v.) This allows us to identify which vertices are good without exposing actual locations of arcs. Indeed, given $d_i^-(v)$ and $d_i^+(v)$, the unexposed endvertices of the $d_i^+(v)$ out-arcs and $d_i^-(v)$ in-arcs visible from v are two subsets of $B_i \setminus \{v\}$ of sizes $d_i^+(v)$ and $d_i^-(v)$ chosen independently and u.a.r. (and also independently of the arcs that are visible from other vertices).

For every vertex v that is not good, we expose all arcs in $\mathcal{D}_{n,p_-} = \mathcal{D}^{\text{in}}_{n,p'_-} \cup \mathcal{D}^{\text{out}}_{n,p'_-}$ (in- and out-arcs, visible and invisible) that are incident with v, and label these arcs *discovered*. This information determines whether v is bad or dangerous. Hence, all vertices are classified into the good, bad and dangerous types defined in Section 4. We also label every vertex that is not good as *discovered*. (Good vertices remain *undiscovered* for now, but this will change later on, as we reveal more information about $\mathcal{D}_{n,p_-}$ and label some good vertices as discovered.) For each $i \in [k]$, let $\hat{B}_i$ denote the set of undiscovered vertices in bin B_i. Note that some of the arcs that are visible from undiscovered (i.e. good) vertices may be discovered, but these are few as we shall see. For each undiscovered vertex v and each bin B_i, let $u_i^+(v)$ be the number of undiscovered out-arcs visible from v and whose other end is in B_i, and let $U_i^+(v)$ be the set of unexposed endvertices of these out-arcs. Similarly, let $u_i^-(v)$ be the number of undiscovered in-arcs visible from v and whose other end is in B_i, and let $U_i^-(v)$ be the set of unexposed endvertices of these in-arcs. Property **H1** of $\mathcal{H}$ guarantees the following:

U1: For any bins B_j, B_i (possibly with $i = j$) and any undiscovered vertex $v \in B_j$,

$$u_i^+(v) \ge d_i^+(v) - 4k \ge 2, \qquad u_i^-(v) \ge d_i^-(v) - 4k \ge 2.$$

Moreover, since the unexposed endpoint of each undiscovered arc remains random, we get:

U2: $U_i^+(v)$ and $U_i^-(v)$ are two subsets of $\hat{B}_i \setminus \{v\}$ of sizes $u_i^+(v)$ and $u_i^-(v)$ chosen independently and u.a.r. (and also independently of the arcs that are visible from other vertices).

Note that **U1**–**U2** are desired properties in light of Lemma 8. We shall see that they remain valid after we expose additional arcs and slightly modify our digraph.

Unfortunately, the arcs in $\mathcal{D}_{n,p_-}$ do not suffice to build a π-HC, since the necessary conditions of event $\mathcal{A}$ do not hold at time p_- (in view of our assumption that $p_- < p_*$). To that effect, we expose all the arcs in $\mathcal{D}_{n,p_*}$ that are incident with a dangerous vertex and

label them discovered, as well. Since p_* is a random time, here is a more careful description of this operation. Let us first expose the number r of arcs (but not their locations) in $\mathcal{D}_{n,p_+} \setminus \mathcal{D}_{n,p_-}$, and list them as $e_1, \dots, e_r$ in the order they appear in the process. For each $1 \leq i \leq r$, we determine whether or not e_i is incident with a dangerous vertex, and if it is then we fully disclose the two endvertices of e_i and label e_i discovered. Otherwise, the endpoints of e_i remain unknown (such e_i will not play any role in our construction of the π-HC). Let r_* be the smallest $i \in \{0, 1, \dots, r\}$ such that $\mathcal{D}_{n,p_-} \cup \{e_1, \dots, e_i\}$ satisfies the conditions in event $\mathcal{A}$. From our assumption that $p_- < p_* < p_+$, we have $1 \leq r_* \leq r-1$ and $\mathcal{D}_{n,p_*} = \mathcal{D}_{n,p_-} \cup \{e_1, \dots, e_{r_*}\}$. Let $\mathcal{D}_* = \mathcal{D}_{n,p_-} \cup \{e_i : e_i \text{ is discovered}, 1 \leq i \leq r_*\}$. By construction $\mathcal{D}_*$ satisfies $\mathcal{A}$ and $\mathcal{H}$ (since $\mathcal{D}_* \subseteq \mathcal{D}_{n,p_*} \subseteq \mathcal{D}_{n,p_+}$ and $\mathcal{H}$ is monotone decreasing). Moreover, the additional newly discovered arcs in $\{e_1, \dots, e_{r_*}\}$ are disjoint with the undiscovered arcs of $\mathcal{D}_{n,p_-}$, since the latter are only incident with good vertices. Hence, the joint distribution of the sets $U_i^{\pm}(v)$ of unexposed endpoints of undiscovered arcs is unaffected by the additional information revealed, and in particular properties **U1**–**U2** are still valid. Our goal is to build a π-HC in $\mathcal{D}_*$ (or just a $(\rightarrow, \leftarrow)$-HC when $\pi = (\rightarrow, \leftarrow, \rightarrow, \leftarrow)$ and $n \equiv 2 \mod 4$). In the sequel, we will expose some additional information about $\mathcal{D}_*$ and make slight modifications to it (such as moving some vertices from one bin to another). By abusing slightly the notation, we will still use $\mathcal{D}_*$ to denote the probability distribution of the modified random digraph given all the exposed information.

Step 2 – short path collection: Our next step is to build a pairwise vertex-disjoint collection $\mathscr{P}$ of short paths, each of which contains exactly one vertex that is not good. To make this more precise, we first consider the case where $k \mid n$ (i.e. either π is non-alternating or $\pi = (\rightarrow, \leftarrow, \rightarrow, \leftarrow)$ with $4 \mid n$). For each non-good vertex w (that is, w is bad or dangerous), we will build a path $P(w) = v_1 v_2 \cdots v_{6k+1}$ of length $6k$ with the following properties: w is an internal vertex of $P(w)$; all the vertices in $P(w)$ are good except for w (so in particular $w \mapsto P(w)$ is a bijection between non-good vertices and paths in $\mathscr{P}$); each arc $v_i v_{i+1}$ is oriented as π_i for $i = 1, \dots, 6k$; and each $v_i \in B_i$ for $i = 1, \dots, 6k+1$. (Recall that modular arithmetic is used for both bins and pattern positions so, for instance, $v_{6k+1} \in B_1$.) Note that in order to achieve this last property, we may have to move some vertices from one bin to another, but this will be done while keeping the bin sizes unchanged. Moreover, only internal vertices of paths in $\mathscr{P}$ may be moved to a different bin. When $\pi = (\rightarrow, \leftarrow, \rightarrow, \leftarrow)$ with $n \equiv 2 \mod 4$, we define the collection of paths $\mathscr{P}$ as above with the only proviso that exactly one of the paths in $\mathscr{P}$ must have length $6k+2$ instead of $6k$. In this longer path $P = v_1 v_2 \cdots v_{6k+3}$, each vertex $v_i \in B_i$ for $i = 1, \dots, 6k+2$ but $B_{6k+3} \in B_1$ (so the path cycles 6 times through bins B_1, B_2, B_3, B_4, and then visits B_1, B_2 and again B_1), and arcs are oriented in an alternating fashion with $v_i v_{i+1}$ oriented as π_i for $i = 1, \dots, 6k+2$.

We will show that we can build a such a collection of paths $\mathscr{P}$. Assume inductively that we have already created a sub-collection $\mathscr{P}_0 \subset \mathscr{P}$ of such paths (possibly $\mathscr{P}_0 = \emptyset$ is the trivial sub-collection) and let w_2 be a non-good vertex not contained in any path in $\mathscr{P}_0$. Our goal is to build a new path $P(w_2)$ with all the desired properties and add it to $\mathscr{P}_0$. We will first suppose that $k \mid n$ (that includes the case when $\pi = (\rightarrow, \leftarrow, \rightarrow, \leftarrow)$ with $4 \mid n$), and then show how to modify the construction if $\pi = (\rightarrow, \leftarrow, \rightarrow, \leftarrow)$ and $n \equiv 2 \mod 4$.

Step 2a – starting path $P(w_2)$: Recall that $N(w_2) = N_{\mathcal{D}_*}(w_2)$ is the set of neighbours of w_2 (that is, those vertices that are joined to w_2 by arcs in $\mathcal{D}_*$ with any label and orientation). This set has already been exposed since w_2 is not good. We claim that we can pick two different neighbours $w_1, w_3 \in N(w_2)$ that are good and are not contained in any path in $\mathscr{P}_0$. Moreover, for the alternating pattern, we can pick w_1, w_3 so that arcs w_1w_2 and w_3w_2 are either both oriented out of w_2 or both into w_2. Indeed, if w_2 is bad, then $|N(w_2)| \geq |N_{\mathcal{D}_{n,p_-}}(w_2)| \geq 4k+3$ by the definition of being bad. Then $N(w_2)$ contains at least three good vertices not contained in any path in $\mathscr{P}_0$ by property **H5** of Lemma 12. By the pigeonhole principle, we can pick two such vertices w_1 and w_3 that are joined to w_2 by arcs with the same orientation. On the other hand, if w_2 is dangerous, all its neighbours are good and do not intersect any path in $\mathscr{P}_0$ by property **H4**. In that case, event $\mathcal{A}$ guarantees the existence of appropriate vertices w_1, w_3 (with arcs w_1w_2 and w_3w_2 both oriented out of w_2 or both into w_2, in the alternating case). Note that this is the only point in the construction of path $P(w_2)$ that we may use arcs in $\mathcal{D}_*$ that are not in $\mathcal{D}_{n,p_-}$.

Step 2b – growing path $P(w_2)$: We first show that we can pick a position $j \in [k]$ in the pattern such that w_1w_2 and w_2w_3 are oriented as π_j and π_{j+1}. First suppose that arcs w_1w_2 and w_3w_2 are both oriented out of w_2 or both into w_2 (note that these are the only two possible cases in the alternating case, from our choice of w_1 and w_3). Clearly for $\pi = (\rightarrow, \leftarrow, \rightarrow, \leftarrow)$ and also for any primitive pattern π of length $k \geq 3$, there must be some $j \in [k]$ such that $\pi_j = \leftarrow$ and $\pi_{j+1} = \rightarrow$ (and similarly some $j' \in [k]$ such that $\pi_{j'} = \rightarrow$ and $\pi_{j'+1} = \leftarrow$). In either case, we can pick a position j with w_1w_2 and w_2w_3 oriented as π_j and π_{j+1}, as claimed. For $k \geq 3$, there is an additional case to be considered. Suppose that exactly one of the arcs w_1w_2 and w_3w_2 is oriented out of w_2 and the other one is into w_2. Since π is not the alternating pattern, there must be some $j \in [k]$ such that $\pi_j = \pi_{j+1} = \rightarrow$ or $\pi_j = \pi_{j+1} = \leftarrow$. Hence (switching vertices w_1 and w_3 if needed), we can conclude that w_1w_2 and w_2w_3 are oriented as π_j and π_{j+1}, also for this case. Let $B_{a_1}, B_{a_2}, B_{a_3}$ (possibly $a_1 = a_2 = a_3$) be the bins containing vertices w_1, w_2, w_3, respectively. We will find a path $P(w_2) = w_{-k-j+2} \cdots w_0w_1w_2w_3 \cdots w_{5k-j+2}$ of length $6k$ in $\mathcal{D}_*$ with the following properties: $P(w_2)$ does not intersect any path in $\mathscr{P}_0$; all vertices $w_{-k-j+2}, \ldots, w_{5k-j+2}$ (of course, except for w_2) are good; each arc w_iw_{i+1} is oriented as π_{i+j-1} for $i = -k-j+2, \ldots, 5k-j+1$; and each $w_i \in B_{i+j-1}$ for $i = -k-j+2, \ldots, 5k-j+2$, except for

$$\begin{cases} w_1 \in B_{a_1} \\ w_2 \in B_{a_2} \\ w_3 \in B_{a_3} \end{cases} \qquad \text{and} \qquad \begin{cases} w_{a_1+2k-j+1} \in B_j \\ w_{a_2+3k-j+1} \in B_{j+1} \\ w_{a_3+4k-j+1} \in B_{j+2}. \end{cases} \tag{2}$$

The bins of vertices w_i and $w_{a_i+(i+1)k-j+1}$ will be swapped for $i = 1, 2, 3$ (unless $a_i = i+j-1$, in which case both w_i and $w_{a_i+(i+1)k-j+1}$ will be placed correctly in the same bin $B_{a_i} = B_{i+j-1}$). Note that since $j, a_1, a_2, a_3 \in [k]$ and $k \geq 3$,

$$-k-j+2 < 1 < 2 < 3 < a_1+2k-j+1 < a_2+3k-j+1 < a_3+4k-j+1 < 5k-j+2,$$

so in particular all six vertices $w_1, w_2, w_3, w_{a_1+2k-j+1}, w_{a_2+3k-j+1}, w_{a_3+4k-j+1}$ must be distinct and internal vertices of the path. For each $i = -k-j+2, \ldots, 5k-j+2$, let B_{α_i} be the

bin where we wish to find vertex w_i. Then $\alpha_i = i + j - 1$, with the possible exceptions described in (2). We will grow our path $P(w_2)$ by adding vertices to each end of $w_1w_2w_3$, one vertex at a time. We can achieve this as follows. For each $i = 3, \dots, 5k - j + 1$, expose all in- and out-arcs of $\mathcal{D}_*$ incident with w_i (visible and invisible from w_i). We label all these exposed arcs as discovered and vertex w_i is labelled discovered as well. Since w_i is good, it has at least $4k + 2$ neighbours w in bin $B_{\alpha_{i+1}}$ with w_iw oriented as π_{i+j-1}. By property **H5**, at least 2 of these vertices are good and not in any path $P \in \mathscr{P}_0 \cup \{w_1w_2w_3 \cdots w_i\}$. Pick one of them and call it w_{i+1}. By an analogous argument, we also grow the path from the other end, by adding a suitable vertex w_{i-1} for each $i = 1, 0, -1, \dots, -k - j + 3$, and thus complete our path $P(w_2)$. This path has all the required properties, except for the fact that vertices listed in (2) may be placed in the wrong bin (depending on the values of a_1, a_2, a_3). Hence, we swap the bins of the vertices in the path that were misplaced: that is, w_i is moved to bin B_{i+j-1} and $w_{a_i+(i+1)k-j+1}$ is moved to B_{a_i}, for $i = 1, 2, 3$. Note that the sizes of the bins remain unchanged and thus balanced. Moreover, writing $P(w_2) = v_1v_2 \cdots v_{6k+1}$ with $v_i = w_{i-k-j+1}$, we have that every vertex v_i ($i = 1, \dots, 6k + 1$) belongs to bin B_i, and each arc v_iv_{i+1} ($i = 1, \dots, 6k$) is oriented as π_i, as required. Then, we can extend $\mathscr{P}_0$ to $\mathscr{P}_0 \cup \{P(w_2)\}$ and inductively obtain our desired collection of paths $\mathscr{P}$.

When $\pi = (\rightarrow, \leftarrow, \rightarrow, \leftarrow)$ and $n \equiv 2 \mod 4$, we just need to extend one of the paths $P = v_1v_2 \cdots v_{6k+1}$ in $\mathscr{P}$ two more steps. When doing so, we must make sure that the two additional vertices v_{6k+2}, v_{6k+3} belong to the right bins and the edges $v_{6k+1}v_{6k+2}, v_{6k+2}v_{6k+3}$ have the appropriate orientations. This can be achieved by the exact same argument as above, and thus we omit the details. Note that we have tacitly assumed that $\mathscr{P}$ contains at least one path, which would not be true if all the vertices of $\mathcal{D}_*$ were good[1]. In that case, we could still create one path by applying Step 2a to a good vertex w_2, and extending this path as in Step 2b.

Step 3 – path contraction: A crucial property in the above construction of $\mathscr{P}$ is that we have only exposed (and labelled discovered) those arcs that are incident with discovered vertices, which are precisely the internal vertices of the paths in $\mathscr{P}$. All the other vertices remain undiscovered (and are good). Note that some arcs incident with undiscovered edges may be discovered. However, in view of **H5** (applying Lemma 12 with $\mathscr{P}_0 = \mathscr{P}$), we infer that each undiscovered vertex v is incident with at most $4k$ discovered arcs. Hence, property **U1** remains valid for any undiscovered vertex $v \in B_j$ and bins B_i, B_j, with the updated values of $\hat{B}_i$, $U_i^{\pm}(v)$ and $u_i^{\pm}(v)$. Moreover, since we did not reveal any information about undiscovered arcs, the unexposed endpoints of the undiscovered arcs visible from v remain random, and therefore property **U2** also holds. Now we contract each path in $\mathscr{P}$ into a vertex. Vertices obtained from a contracted path are called *fat* and are placed in bin B_1, while other vertices are called *ordinary*. We declare fat vertices to be undiscovered, so all vertices in the contracted graph are undiscovered. Let n' be the number of vertices (ordinary or fat) that remain in the digraph after the path contractions. Clearly $n/(7k) \le n' \le n$, since each path has fewer than $7k$ vertices. In particular, $n' \to \infty$. We claim that $k \mid n'$ and moreover, after contracting the paths, every bin contains exactly n'/k vertices. Indeed, when $k \mid n$, each path in $\mathscr{P}$ has 6 vertices in each bin B_i ($i = 2, \dots, k$) and 7 vertices in

[1]It is easy to show that a.a.s. this does not happen, but we do not make use of this fact.

bin B_1, so path contractions maintain the bin sizes balanced. On the other hand, when $\pi = (\rightarrow, \leftarrow, \rightarrow, \leftarrow)$ and $n \equiv 2 \mod 4$, two additional vertices—one in bin B_1 and one in B_2—are contracted (as a result of the longer path), so bins B_1, B_2, B_3, B_4 become balanced as well in this case.

Step 4 – end of the proof: At this final stage, we will only need arcs between vertices in consecutive bins B_i and B_{i+1}, for each $i \in [k]$, so we delete all the remaining arcs. Moreover, for each fat vertex v obtained by contracting a path $P = v_1 \cdots v_{6k+1}$, we delete all arcs (in- and out-arcs, visible and invisible) that are incident with v_1 except for those whose other end is in bin B_k, which remain unexposed. Similarly, we delete all arcs that are incident with v_{6k+1} except for those whose other end is in bin B_2. In other words, fat vertex v plays the role of vertex v_1 (respectively, v_{6k+1}) when concerned with arcs between bins B_1 and B_k (respectively, B_1 and B_2). Let us call the resulting digraph $\mathcal{D}_{**}$. (Recall that $\mathcal{D}_{**}$ has $n' \to \infty$ vertices with $k \mid n'$, and each bin contains exactly n'/k vertices.) We claim that, after all these edge deletions, properties **U1**–**U2** remain valid for any vertex $v \in B_j$ and bin B_i with $i \in \{j-1, j+1\}$.[2] This is clearly true if v is a fat vertex corresponding to a path $P = v_1 \cdots v_{6k+1}$, since v inherits the arcs visible from v_1 with opposite end in B_k and the arcs visible from v_{6k+1} with opposite end in B_2. On the other hand, for any ordinary vertex $v \in B_j$ and bin B_i ($i \in \{j-1, j+1\}$), the only arcs visible from v that may have been deleted are those incident with a fat vertex (and thus with some vertex of a path in $\mathscr{P}$). Hence, by property **H5** as before, we conclude that **U1** remains valid. Property **U2** also follows from the fact that we have not exposed any additional information about the surviving arcs and thus they remain uniformly distributed. Finally, **U1** and **U2** imply that the contracted digraph $\mathcal{D}_{**}$ contains a copy of $D_{S\text{-in},T\text{-out}}$ (on n' vertices with $n' \to \infty$) with $s_{i,i+1}, s_{i,i-1}, t_{i,i-1}, t_{i,i-1} \geq 2$. So we can apply Lemma 8 and find a π-HC in $\mathcal{D}_{**}$ a.a.s. Finally, when π is primitive of length $k \geq 3$, un-contracting the paths yields the desired π-HC in $\mathcal{D}_* \subseteq \mathcal{D}_{n,p_*}$. When $\pi = (\rightarrow, \leftarrow, \rightarrow, \leftarrow)$, we simply get a $(\rightarrow, \leftarrow)$-HC. This finishes the proof of Theorem 2. Corollaries 3 and 5 follow immediately as a consequence of Lemma 10.

6 Final Comments

We have proved a hitting time result for the existence of a Hamilton cycle with edges oriented according to a periodic pattern. We have shown that in some cases this typically requires asymptotically one half of that needed for a consistently oriented Hamilton cycle. It would be of some interest to discover how many such cycles there are typically and as to whether there are all such cycles, up to a bound on the length of the pattern π. As already mentioned, Ferber and Long [4] have shown that Hamilton cycles with arbitrary orientations occur a.a.s. at $m \approx n \log n$. It would be of interest to understand which class of patterns occur a.a.s. for $m \leq cn \log n$ where $0 < c < 1$ is a constant.

[2]The proof of this claim makes use of the fact that $k \geq 3$ and hence arcs between bins B_1 and B_2 are different from those between B_1 and B_k. This justifies our choice to use the pattern $(\rightarrow, \leftarrow, \rightarrow, \leftarrow)$ in the analysis of the alternating case. Other approaches are also possible, but we believe this is the simplest one.

References

[1] M. Ajtai, J. Komlós and E. Szemerédi. The first occurrence of Hamilton cycles in random graphs, *Annals of Discrete Mathematics* 27 (1985) 173-178.

[2] C. Cooper and A.M. Frieze, The limiting probability that a-in, b-out is strongly connected, *Journal of Combinatorial Theory* B 48 (1990) 117-134.

[3] C. Cooper and A.M. Frieze, Hamilton cycles in random graphs and directed graphs, *Random Structures and Algorithms* 16 (2000) 369-401.

[4] A. Ferber and E. Long, Packing and counting arbitrary Hamilton cycles in random digraphs, *Random Structures and Algorithms* 54 (2019) 499-514.

[5] A.M. Frieze, An algorithm for finding hamilton cycles in random digraphs, *Journal of Algorithms* 9 (1988) 181-204.

[6] A.M. Frieze, Hamilton Cycles in Random Graphs: a bibliography.

[7] A.M. Frieze and M. Karoński, Introduction to Random Graphs, Cambridge University Press, 2015.

[8] R. Montgomery, Universality for spanning cycles in random directed graphs.

[9] D.W. Walkup, Matchings in random regular bipartite graphs, *Discrete Mathematics* 31 (1980) 59-64.